A Hybrid Queue-based Bayesian Network Framework for Passenger Facilitation Modelling

Paul Pao-Yen Wu[1], Jegar Pitchforth[2], Kerrie Mengersen[3]

Queensland University of Technology, Mathematical Sciences School, 2 George Street, Brisbane, Australia. Phone: 61-7-31389828, Fax: 61-7-31381827.

Abstract

This paper presents a novel framework for the modelling of passenger facilitation in a complex environment. The research is motivated by the challenges in the airport complex system, where there are multiple stakeholders, differing operational objectives and complex interactions and interdependencies between different parts of the airport system. Traditional methods for airport terminal modelling do not explicitly address the need for understanding causal relationships in a dynamic environment. Additionally, existing Bayesian Network (BN) models, which provide a means for capturing causal relationships, only present a static snapshot of a system.

A method to integrate a BN complex systems model with stochastic queuing theory is developed based on the properties of the Poisson and Exponential distributions. The resultant Hybrid Queue-based Bayesian Network (HQBN) framework enables the simulation of arbitrary factors, their relationships, and their effects on passenger flow and vice versa.

A case study implementation of the framework is demonstrated on the inbound passenger facilitation process at Brisbane International Airport. The predicted outputs of the model, in terms of cumulative passenger flow at intermediary and end points in the inbound process, are found to have an R^2 goodness of fit of 0.9994 and 0.9982 respectively over a 10 hour test period. The utility of the framework is demonstrated on a number of usage scenarios including causal analysis and 'what-if' analysis. This framework provides the ability to analyse and simulate a dynamic complex system, and can be applied to other socio-technical systems such as hospitals.

Keywords: complex systems, dynamic system, modelling, Bayesian Network, airport, passenger facilitation

---

1. Introduction

Modern airports face a number of challenges including growing passenger traffic and its effects on capacity and congestion as well as regulatory and market changes such as the proliferation of Low Cost Carriers (LCC)

URL: p.wu@qut.edu.au (Paul Pao-Yen Wu), jegar.pitchforth@qut.edu.au (Jegar Pitchforth), k.mengersen@qut.edu.au (Kerrie Mengersen)

[1]Corresponding author.

[2]Author.

[3]Author.

(de Neufville and Odoni, 2003; Nombela et al., 2004). In addition, there is a significant time and monetary cost associated with the construction or renovation of infrastructure needed for airport operations (Odoni and de Neufville, 1992). The challenges are further compounded by the fact that the airport is a complex system. There are multiple stakeholders who at times have conflicting objectives (Eilon and Mathewson, 1973; Schultz and Fricke, 2011). Additionally, there are complex interactions and interdependencies between stakeholders and between different parts of the airport system (Zografos and Madas, 2006; Manataki and Zografos, 2009). Finally, the system itself is highly dynamic and there is significant uncertainty on its performance (Manataki and Zografos, 2009; Lui et al., 1972).

From an operational perspective, a model can integrate the diverse elements of the airport complex system to help the user understand how the airport is likely to perform under different operational scenarios. This can be invaluable in supporting capacity planning, operational planning and design, and airport performance review (Wu and Mengersen, 2013). Examples of this in the past include the application of stock and flow modelling to ascertain whether the terminal infrastructure was sufficient to handle the increased traffic related to the Athens Olympic Games (Zografos and Madas, 2006). Additionally, Jim and Chang (1998) discuss how simulation modelling can specifically assist the final design stage of terminal development. Brunetta et al. (1999) describe another example of how modelling is used to assist capacity estimation at Milan Linate and Malpensa 2000 airports.

The efficient flow of passengers and/or aircraft is a key goal in the operations management of an airport (de Neufville and Odoni, 2003). However, this is a challenging task because the cause of long queues and other problems affecting flow are not necessarily obvious due to complex interactions and interdependencies within the airport complex system. A number of models have been presented that provide a means to indirectly determine the cause of flow performance through simulation, such as (Manataki and Zografos, 2009; Wilson et al., 2006; Eilon and Mathewson, 1973). This paper proposes a modelling framework that can explicitly capture the flow of passengers and its relationship to the different factors affecting its performance, thus enabling decision makers to focus on the root cause of performance issues. The work is demonstrated on a case study application, which is the inbound passenger facilitation process.

1.1. The Passenger Facilitation Modelling Problem

Consider the passenger terminal, which is one of the key subsystems within the airport environment (de Neufville and Odoni, 2003). This paper seeks to address the challenges of modelling the inbound passenger facilitation process as defined by the Passenger Facilitation Taskforce (2009) to provide decision support. Annex 9 of the Convention on International Civil Aviation provides standards and recommended practices for passenger facilitation, which is the process that "assists the free flow of passengers and goods across the border whilst upholding border integrity and/or sovereignty" (International Civil Aviation Organisation, July, 2005). The selected case study focuses on arriving passengers (i.e. inbound passengers) and was undertaken as part of the Airports of the Future (AotF) project and involved government and industry partners, including the Australian Customs and Border Protection Service (Customs and Border Protection), Brisbane Airport Corporation, and the Department of Agriculture, Fisheries and Forestry (DAFF) Biosecurity.

The work was initially motivated by the requirements of the National Passenger Facilitation Committee (NPFC) performance framework initiative, which sought to establish standardised levels of facilitation performance such as in terms of passenger wait time, congestion and throughput across Australian international airports. Note that these are all measures of passenger flow, which underpin the passenger facilitation process. Therefore a model of inbound passenger facilitation must capture **passenger flow** explicitly. Note also that the spatial aspect of passenger movement and the spatial constraints of the enviornment also play a role in passenger flow, hence **space** and/or the effects of space also need to be captured.

However, engagement with stakeholders revealed that there were other operational factors and objectives such as border security and biosecurity that also needed to be met. Passenger flow can affect the performance of factors such as border risk and biosecurity risk and vice versa; for example, increased security procedures lead to lower risk but increased queuing times (Wilson et al., 2006).

The stakeholders also indicated a need for explicit, quantitative causal analysis; i.e. the ability to quantitatively characterise the relationship between factors such as congestion, passenger demographics (e.g. age, nationality), and 'performance' factors such as passenger processing time and throughput. Eilon and Mathewson (1973) for example use a regression sub-model within an simulation model to mathematically capture the relationship between congestion (number of passengers) and passenger delay time. Such a sub-model enables decision makers to better understand how one factor affects another, and hence ascertain the root cause of performance issues. As a result, the model also needs to be **extensible to other factors** and also enable **explicit, quantitative causal analysis**.

In addition, due to the variability in day-to-day operations, such as due to weather, flight delays or equipment malfunction, the stakeholders require the capability to simulate **'what-if'** scenarios. Note that the implementation of any model requires a means for updating or learning model parameters. This is especially the case for an airport as it is a constantly changing environment (de Neufville and Odoni, 2003). In summary, the main modelling requirements are: (i) to capture passenger flow, (ii) incorporate the effects of space, (iii) be extensible to other factors, (iv) enable explicit, quantitative causal analysis, and (v) enable 'what-if' analysis.

1.2. Summary

This paper presents a novel Hybrid Queue-based Bayesian Network (HQBN) framework for modelling passenger facilitation. The framework integrates a Bayesian Network (BN) model of the passenger facilitation system with a stochastic queuing model of passenger flow based on the Poisson process. Using the proposed framework, it is possible to leverage the inherent explicit, quantitative causal analytic capabilities of the BN to capture the relationships between passenger flow and the various factors that make up the airport terminal system. A review of the existing literature by Wu and Mengersen (2013) has revealed that existing work does not simultaneously address all of the requirements identified in Section 1.1, especially the combination of passenger flow modelling and explicit, quantitative causal analysis.

The paper is organised as follows. Section 2 reviews the existing terminal modelling literature with respect to papers on inbound or outbound passenger facilitation. Based on these findings, the proposed framework is presented in section 3. This method is demonstrated in section 4 for a case study on the inbound passenger facilitation process that was performed at Brisbane International Airport. The findings are discussed in section 4.3 and conclusions drawn in section 5.

2. Existing Work

Existing models of the airport passenger terminal that are capable of simulating and analysing passenger flows are predominantly used for operational planning and design (Wu and Mengersen, 2013). This section provides a brief review of such models for inbound and/or outbound passenger facilitation.

Many existing operational planning models adopt an Agent Based Modelling (ABM) approach (Wu and Mengersen, 2013). These models simulate the behaviour and movement of individual passengers (i.e. the 'agents') given the spatial layout of the terminal building as demonstrated for example by (Schultz and Fricke, 2011; Kleinschmidt et al., 2011; Takakuwa and Oyama, 2003; Kiran et al., 2000). At an individual passenger level, the movement of the passengers is captured using methods such as the social force model where passengers are attracted to a destination and repelled by obstacles and other passengers (Helbing and Molnar, 1995). The agents are assigned 'goals' according to the steps in the passenger facilitation process.

In view of the modelling requirements described in section 1.1, ABMs capture both passenger flow and space. Additionally, they can be extended to model other factors such as security risk (Wilson et al., 2006; Koch, 2004). Moreover, passenger facilitation ABMs have been applied to a range of 'what-if' scenarios relating to different (or new) airport configurations (Takakuwa and Oyama, 2003; Wilson et al., 2006), different flight schedules (Jim and Chang, 1998) and different resource assignment schedules (Eilon and Mathewson, 1973).

However, ABMs do not provide an explicit representation or quantification of the causal relationships between different elements within the model. These relationships are established indirectly through simulation; hence, it is necessary to apply another modelling method to obtain the explicit, quantitative, cause-and-effect relationships. Eilon and Mathewson (1973) for example achieve this with a regression sub-model to understand the cause-and-effect relationship between congestion and passenger delays; Appelt et al. (2007) use hypothesis testing. A summary of the preceding discussion on ABM with respect to the modelling requirements in section 1.1 is provided in Table 1.

An alternate approach to modelling passenger facilitation is that of queuing theory and queue networks. Examples include deterministic models like that presented by Newell (1971) and Brunetta et al. (1999), and stochastic models like that presented by Bevilacqua and Ciarapica (2010). Such models inherently provide a means to capture passenger flow as they are predicated on passenger throughput and its relationship to time, namely, waiting time and processing time (Tosic, 1992). Yanagisawa et al. (2013) present a queuing theory based approach to the detailed analysis of queue performance where spatial effects are captured based on their impact on temporal performance. While these models are similar to ABMs in that both are simulation models capable of 'what-if' analysis, queuing models such as

Table 1: Summary of existing passenger facilitation models and their features as identified in Section 1.1. Note that ✓ denotes that the requirement is met, and – shows the requirement is not directly/explicitly met.

	Modelling Requirements				
Method	Passenger Flow	Extensible to Other Factors	Explicit, Quantitative Causal Analysis	Space	What-if Analysis
ABM	✓	✓	–	✓	✓
Queuing theory	✓	–	✓	–	✓
System dynamics	✓	✓	–	–	✓
HQBN (proposed method)	✓	✓	✓	–	✓

(Bevilacqua and Ciarapica, 2010; Newell, 1971; Brunetta et al., 1999; Tosic, 1992) tend to be focused purely on the passenger flow aspect and do not incorporate other factors such as security risk.

Unlike ABMs, space is represented implicitly through time durations for passenger facilitation activities. Additionally, unlike ABMs, queuing theory presents an analytical means for relating passenger flow to the number of processing servers and processing time (Ross, 2010). A summary of passenger facilitation queuing models with respect to the modelling requirements is presented in Table 1.

Manataki and Zografos (2009, 2010) present an alternative approach based on system dynamics. The presented model is also a simulation model, however, it simulates passenger flows at a population level rather than individual interactions like an ABM. Like ABMs, it is possible to perform 'what-if' analysis; in addition, other factors can be captured as demonstrated by the example presented for security screening (Manataki and Zografos, 2009). Although explicit causal analysis is not presented by Manataki and Zografos (2009, 2010), it is possible to adopt a similar approach to Eilon and Mathewson (1973) where a statistical model is applied to simulation results to explicitly and quantitatively represent cause-effect relationships. Finally, unlike ABMs, this approach does not capture space explicitly. A summary of the system dynamics approach compared against the modelling requirements is presented in Table 1.

Bayesian Networks (BNs) provide one approach for addressing the need for explicitly and quantitatively capturing cause-effect relationships between the factors that make up a system (Pearl, 1988). Each relationship is quantified with a Conditional Probability Table (CPT), which provides an additional capability to capture uncertainty in the airport system. Such uncertainty is represented in the form of a probability distribution, which can be used to assist decision makers in evaluating not just the expected performance outcome, but also other possible outcomes based on their probability of occurrence. The method has been applied in a diverse range of fields to describe and analyse a wide variety of complex systems. Heckerman et al. (1995) provide a good overview of early applications of BN models; Pourret et al. (2008) provide a more recent survey.

A BN, also referred to as a belief network, Bayesian belief network or inference diagram, provides a visual representation of the factors and the relationships between them for the system being modelled (Heckerman et al., 1995). Such a visual representation greatly aids communication and understanding. Behind each node in the graph, however, is a CPT that quantifies the relationships that are presented visually.

However, the airport terminal is a highly dynamic environment and a BN only provides a static snapshot of the problem domain (Kjrulff, 1995). Dynamic BNs (DBNs) are an extension of BNs into the time domain based on the Markovian assumption (i.e. the current time step is only affected by the preceding time step) (Kjrulff, 1995). The DBN presents a potential solution to the airport terminal modelling problem, however, the challenge remains of how to represent and simulate the relationship between passenger flow and passenger cycle or dwell time.

It can be seen from Table 1 that simultaneously addressing all of the modelling requirements is challenging. A framework that can simultaneously provide explicit, quantitative causal analysis (i.e. quantify cause-effect relationships) while being extensible to different operational factors has remained a challenge in the passenger facilitation modelling space. The next section presents a novel method to address this gap in the literature. For a more detailed review of the existing literature, refer to Wu and Mengersen (2013).

3. Hybrid Queue-based Bayesian Network (HQBN)

A Hybrid Queue-based Bayesian Network (HQBN) approach is proposed based on a combination of the BN (Pearl, 1988) and stochastic queuing theory using the Poisson and Exponential distributions (Ross, 2010). The method combines the ability to capture causal relationships between system factors as per a BN and the dynamic movement of passengers as per a queuing model.

Consider the airport terminal system S which is made up of a number of subsystems $S_i \subset S$ where there are L subsystems and $i = 1, ...L$. Let each subsystem S_i be characterised by:

- $n_i(t_k)$, the expected number of passengers in S_i at discrete time slice t_k,
- $n_{ai}(t_k)$, the expected number of passengers entering (or arriving) at S_i at time t_k,
- $n_{di}(t_k)$, the expected number of passengers exiting (or departing) S_i at time t_k, and
- $\tau_i(t_k)$, the average cycle time or dwell time for S_i. In a deterministic system, $n_{di}(t_k) = m/\tau_i(t_k)$ where m is a measure of the number of parallel exit channels from that area.
- In addition, let $\alpha_{i,j}(t_k)$ denote the probability that passengers will flow from subsystem S_i to S_j at time slice t_k.

Given the above formulation, the following sections address the BN and stochastic queuing aspects of the HQBN framework respectively and how they integrate.

3.1. HQBN: BN Component

The Iterative Bayesian Network Development Cycle (IBNDC) (Johnson et al., 2010) provides the basis for the development of a BN model of the airport terminal system S with respect to flow related performance metrics and

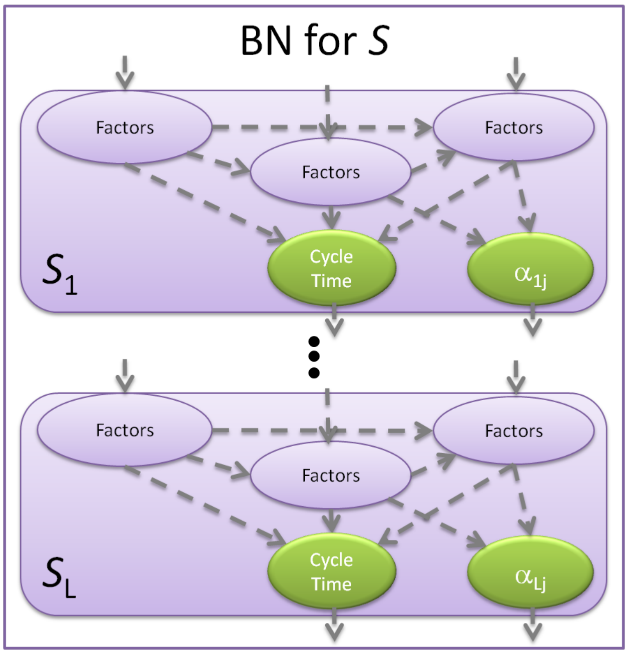

Figure 1: Generic structure for the terminal system BN. Note the interface nodes for τ_i (cycle time) and $\alpha_{i,j}$.

other criteria such as risk and demographics. The BN provides a means to analyse relationships and dependencies such as through forwards or backwards inferencing and sensitivity analysis (Neapolitan, 2004). It also enables simulation of 'what-if' scenarios and provides a means of learning of model parameters (e.g. CPTs) from data (Neapolitan, 2004). Additionally, spatial effects can be captured via factors such as concourse walking time distributions and congestion. Therefore, the BN method can be used to address almost all of the criteria specified in Table **??**. However, the traditional BN does not provide a means to simulate a dynamic system. A DBN can address the dynamic element, however, the challenge remains in establishing the relationships between system factors and passenger flow over time.

Consider a BN model of an airport terminal system S like that described above. Such a BN could, depending on the intended usage scenario of the model, capture such factors and relationships as: passenger demographics (e.g. nationality, age), operational factors (e.g. number of staff, staff experience, process complexity), and other factors (e.g. biosecurity risk, passenger satisfaction) (Wu and Mengersen, 2013). There is no restriction on the number of links (directed edges) or which nodes are linked (e.g. nodes in S_i could influence nodes in S_j), as long as the assumption of a directed acyclic graph is not violated (Pearl, 1988).

According to the formulation, each subsystem S_i has associated with it a cycle time (or dwell time) τ_i and the probability of moving from subsystem S_i to S_j is $\alpha_{i,j}$. Therefore, the BN model of S must include in it, for every subsystem S_i, interface nodes (or factors) corresponding to τ_i and $\alpha_{i,j}$. This is illustrated in Fig. 1. Note that it is assumed that both the cycle time node and $\alpha_{i,j}$ nodes are discrete nodes. These nodes can be discretised to a level as required for the application at hand; for example, where the model is used to assess the likelihood of meeting or exceeding some performance target X, the target itself can be used as a threshold for discretisation (i.e. states of node x are $x < X$ and $x \geq X$).

3.2. HQBN: Stochastic Queuing Component

A complex systems airport terminal model needs to capture passenger flow in terms of time (e.g. dwell time) and movement (i.e. number of passengers), in addition to other factors and perspectives. The preceding section illustrated how a BN can be used to address many of these complex systems modelling requirements as per Table **??**. This section describes a novel method to transform the BN into a dynamic model and make the links between passenger flow and system factors such as demographic and biosecurity factors.

The Poisson and exponential distributions have been used to model many real world processes, especially for queuing systems (Ross, 2010). Let the movement of passengers from one subsystem to another be assumed to be a Poisson process, where individual movements are random and independent of one another. The Poisson distribution, which characterises the number of occurrences of an event, in this case the number of passengers moving from subsystem S_i to S_j, is defined as follows:

$$P(x) = \frac{e^{-\mu}\mu^x}{x!} \tag{1}$$

where $x = 0, 1, 2, ...$ is the number of occurrences and μ is the mean of the distribution. The Exponential distribution has the following definition:

$$P(t) = \lambda e^{-\lambda t} \tag{2}$$

where $t \geq 0$ is the time variable and $\lambda > 0$ is the average rate per unit of time.

The Poisson and exponential distributions are related such that if the Poisson distribution describes the number of occurrences within a given interval of time, then the length of time between occurrences follows an exponential distribution. Therefore, there are, on average, $\mu = \lambda t$ occurrences per t units of time. In other words, these two distributions share a common parameter, namely, the mean rate λ.

Consider the case where $x = 0$, which can be interpreted as the probability that there are no occurrences in t units of time. Equivalently, $x = 0$ can be considered as the probability that the time T until the first occurrence is greater than t. Therefore, substituting $\mu = \lambda t$ and $x = 0$ into (1) gives:

$$P(x = 0) = P(T > t) = e^{-\lambda t} \tag{3}$$

Equivalently,

$$P(T \leq t) = 1 - e^{-\lambda t} \tag{4}$$

Rearranging, it is found that:

$$\lambda = \frac{-log\,(1 - P(T \leq t))}{t} \tag{5}$$

Consider the discretisation of the cycle time node, τ_i. Using (5), it can be seen that τ_i needs to be discretised into two states where one state gives the belief of being below or equal to a specified cycle or dwell time t, and the other state gives the belief of being above t. Note that as $log(0)$ is undefined, it is necessary to approximate $P(T \leq t) = 1 \approx 1 - \epsilon$ where ϵ is a small number. Using this property of the Poisson and exponential distributions, it is possible to transform between cycle time and the average rate at which passengers move from subsystem to subsystem.

Depending on the application, a finer resolution may be required for the cycle time node τ_i. Consider the case where τ_i is an interval node with $Q > 2$ states and each state q is of the form $\left\{t_l^q \leq T < t_u^q\right\}$ where $t_l^q, t_u^q \geq 0, t_u^q > t_l^q, t_l^q, t_u^q \in \mathbb{R}$.

Based on the properties of the exponential distribution described in (3) and (4), it follows that:

$$P(t_l^q \leq T < t_u^q) = e^{-\lambda_q t_l^q} - e^{-\lambda_q t_u^q} \tag{6}$$

As t_l^q, t_u^q and $P(t_l^q \leq T < t_u^q)$ are part of the state definition and posterior belief of the BN node respectively, it is possible to solve for λ_q numerically for each state q. As the proposed queuing model uses the average rate λ, this can be obtained by taking the expectation over all states:

$$\lambda = \sum_q P(t_l^q \leq T < t_u^q)\lambda_q \tag{7}$$

Note that when $q = 2$, taking the expectation is not necessary as $P(T > t) = 1 - P(T \leq t)$, thus giving the same value for λ as per (3) and (4) respectively.

In the airport system, there are often multiple parallel servers (also referred to as channels or counters or modules) within each subsystem. For example, there may be multiple immigration desks processing passengers in parallel. The processing of passengers by each server can be treated as independent Poisson processes. As the sum of N independent Poisson processes is also a Poisson process with a mean rate equal to the sum of the means, the total number of passengers exiting S_i at time t_k can be expressed as follows (Ross, 2010):

$$n_{di}(t_k) = m\lambda(t_k) \tag{8}$$

where m is the number of servers and $\lambda(t_k)$ is found using (5) or (7) at time slice t_k.

Note that when m is set to equal the total number of passengers $n_i(t_k)$ in S_i, this corresponds to a model of a pure delay (the 'infinite' number of servers scenario) (Ross, 2010). For airport processing areas such as check-in, security screening, immigration, boarding and related processing (on inbound and outbound processes), there is a defined queuing process and m is set to the number of servers (de Neufville and Odoni, 2003). However, for discretionary activities (Popovic et al., 2009) and baggage reclaim, there is no clearly defined server; in these areas, m is set to $n_i(t_k)$ to model a pure delay.

For any subsystem S_j, $n_j(t_k)$ is defined recursively as follows:

$$n_j(t_k + 1) = n_j(t_k) + n_{aj}(t_k - 1) - n_{dj}(t_k) \tag{9}$$

Given that $\alpha_{i,j}(t_k)$ defines the probability that passengers move from S_i to S_j at time t_k, substituting into (9) gives:

$$n_j(t_k + 1) = n_j(t_k) + \sum_{i=1}^{L} \alpha_{i,j}(t_k)n_{di}(t_k) - n_{dj}(t_k) \tag{10}$$

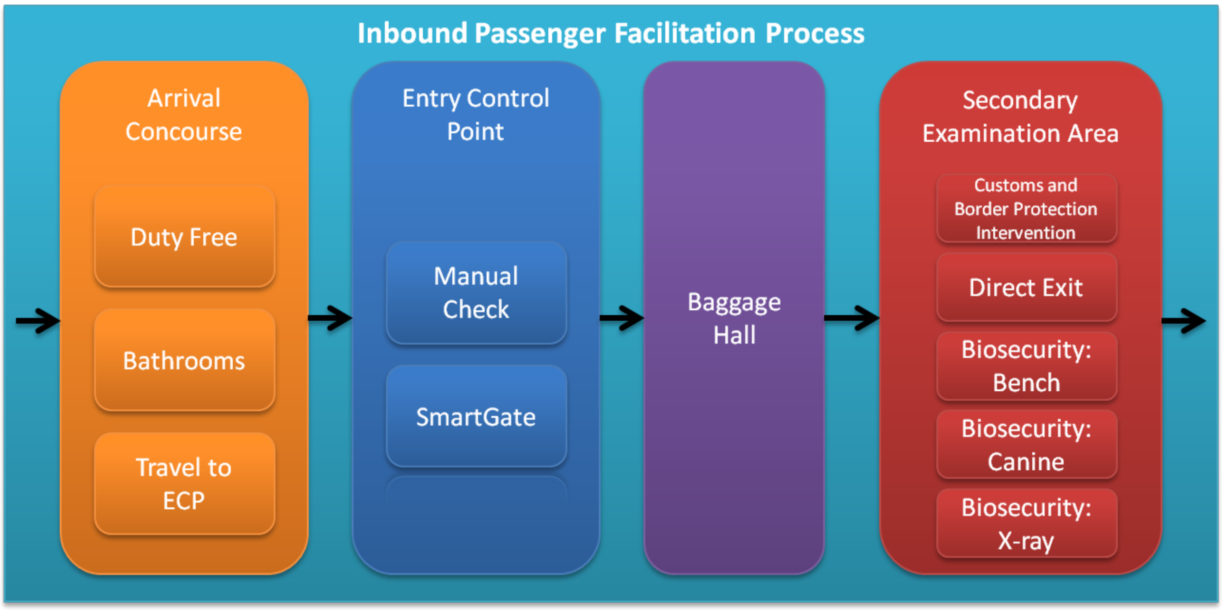

Figure 2: Inbound passenger facilitation process as divided into subsystems.

In summary, the proposed approach simulates over a chosen time period where at each time slice t_k, the inference is performed on the BN component of the model to determine the rates of movement for each area, which then updates the number of passengers in each area for the next time slice t_k+1 using (10). It can be seen that the proposed approach is similar to a DBN (Kjrulff, 1995) in that both are based on discrete time slices under the Markovian assumption, however, the proposed method explicitly deals with the relationship between time and passenger movement.

4. Case Study: Inbound Passenger Facilitation at Brisbane International Airport

This section demonstrates and discusses the implementation of the proposed HQBN framework on the inbound passenger facilitation process at Brisbane International Airport. The scope of the inbound facilitation process is defined as shown in Fig. 2 for incoming passengers arriving from overseas who are entering the country (Passenger Facilitation Taskforce, 2009).

4.1. The Inbound Passenger Facilitation Model

The first step in the development of the Inbound Passenger Facilitation Model (IPFM) was the identification of the system and subsystems. Using Passenger Facilitation Taskforce (2009) as a starting point, the inbound facilitation process was identified through consultation with experts (the airport operator, Customs and Border Protection and Biosecurity) as shown in Fig. 2. It was indicated by these expert end-users that transfer passengers did not need to be included as transfer traffic accounts for less than 1% of international passenger movements at Brisbane (Australian Competition and Consumer Commission (ACCC), 2013). It can be seen that the system comprises elements of processing following aircraft disembarkation and prior to exit from the Secondary Examination Area (SEA). The four main areas of the inbound process are thus: the Arrival Concourse (AC), Entry Control Point (ECP), Baggage Hall (BH) and SEA.

The AC represents a relatively simple functional area with three subsystems:

- Travel to ECP - this subsystem presents a pure delay that captures the distribution of travel times associated with moving from the gate to the ECP.
- Bathrooms - this subsystem captures the delay associated with bathroom usage.
- Duty free - this subsystem captures the delays associated with browsing and purchasing.

All passengers must pass through the 'Travel to ECP' subsystem, however, movement of passengers through the 'Duty Free' and 'Bathrooms' subsystems are discretionary and can occur in any order. Therefore, passengers can transition from any of the three AC subsystems to the ECP. Note that given the scope of the inbound process, it is assumed that all passengers proceed to the ECP (i.e. transit passengers are out of the scope of this model). In this model, the delays are represented as discrete probability distributions based on surveys (e.g. walking time) and/or expert elicitation, and transformed into passenger flow using the proposed HQBN framework (see Section 3.2).

In contrast, the ECP contains only two subsystems where passengers either go to 'Manual Check' or 'SmartGate', but not both. These two subsystems are both queuing systems with multiple service modules. SmartGate has the added complexity of a two-step process whereby passengers must first complete a kiosk step, then a gate (face recognition) step (Passenger Facilitation Taskforce, 2009); in this case, SmartGate could itself be decomposed into two subsystems.

A diverse array of factors that affect ECP performance are captured in the model, reflecting the complexity of the airport system. For example, passenger demographic information, specifically nationality, age and possession of a compatible passport, determines eligibility for SmartGate. The presence or absence of a Customs and Border Protection marshal further affects the probability that an eligible person will use SmartGate. Finally, the flight origin point and the interaction between flights arising from the flight schedule all affect passenger flow in the ECP. Note that many of the factors discussed here also apply in different ways to the other three sub-systems.

The baggage hall represents a simpler subsystem whereby passengers experience a pure delay in waiting for and reclaiming their bags. As passengers exit the ECP, there is a small possibility they will be interviewed by a Biosecurity Officer or Customs and Border Protection Officer or both.

Finally, the last and most complex subsystem is the SEA. Firstly, passengers queue to meet the Customs and Border Protection Officer who acts as a marshal and directs them, according to risk criteria, to direct exit, a Customs and Border Protection intervention, or a Biosecurity intervention. Should the passenger be directed to Biosecurity, they join a further queue to a Biosecurity marshal who directs them onto the various paths that are available.

As a result, the BN for this area needs to capture the variables considered by the respective marshals in order to enable causal analysis of $\alpha_{i,j}(t_k)$; these considerations predominantly revolve around risk, embarkation point of the flight, and passenger declarations (e.g. items being brought into the country). Each of the subsystems of the SEA shown in Fig. 2 can be captured by a queue and service approach as described in section 3.2.

The Key Performance Indicators (KPIs) for the stakeholders involved in the inbound passenger facilitation process include: (i) passenger facilitation and throughput, (ii) dwell time (not more than 45 minutes for the overall process), and (iii) border and biosecurity risk standards[4]. Based on the facilitation process identified above, a HQBN is developed using the framework described in Section 3; the model is tested and validated in Section 4.2 and demonstrated on a number of applications in Section 4.3.

4.2. IPFM Quantification, Testing and Validation

The case study model is quantified with a combination of available data including expert knowledge, immigration data, summary statistics, and Closed Circuit TeleVision (CCTV) based intelligent surveillance methods developed within the AotF project. Where data exists, learning algorithms such as expectation maximisation are applied to learn model parameters (i.e. the CPTs) Neapolitan (2004). Otherwise, expert elicitation is performed to ascertain CPT values Choy et al. (2009). Note that the available data include demographic information and also time stamps of when passengers are cleared through the ECP. Additionally, the intelligent surveillance data comprise timestamped counts of passengers as they move past certain checkpoints, similar to that described by Gongora and Ashfaq (2006).

In terms of available data that can be used to validate model simulated passenger movement, only the time registered count of passengers exiting the ECP can be used as ground truth. This dataset is highly reliable as it is based on the last keystroke before a passenger clears the ECP. On the other hand, the data collected using CCTV intelligent surveillance have inherent errors associated with the visual situation (e.g. occlusions and camera view). The lack of complete data, highlighted by numerous data gaps, demonstrate the need for a complex systems model that can integrate data and knowledge from a variety of information sources.

Passenger flow is often depicted using cumulative ‘arrival’ or ‘departure’ curves that show the accumulated number of passengers who have entered or exited a subsystem over time (de Neufville and Odoni, 2003). These curves have the added advantage that:

1. they show whether the model drifts from the actual count over time,
2. they are not sensitive to fluctuations in the instantaneous entry or exit rate[5],
3. the vertical displacement between successive curves shows the number of passengers in that subsystem,
4. the horizontal displacement between successive curves shows the average dwell time in that subsystem, and
5. they can be used to assess the overall validity of the model as they represent the output of the BN and the model (with respect to passenger flow).

[4]Due to security requirements, these figures can not be divulged.

[5]as the count accumulates (e.g. hundreds or thousands of passengers have exited), a change in the instantaneous rate (typically in the order of tens of passengers) does not produce a big change in the accumulated count.

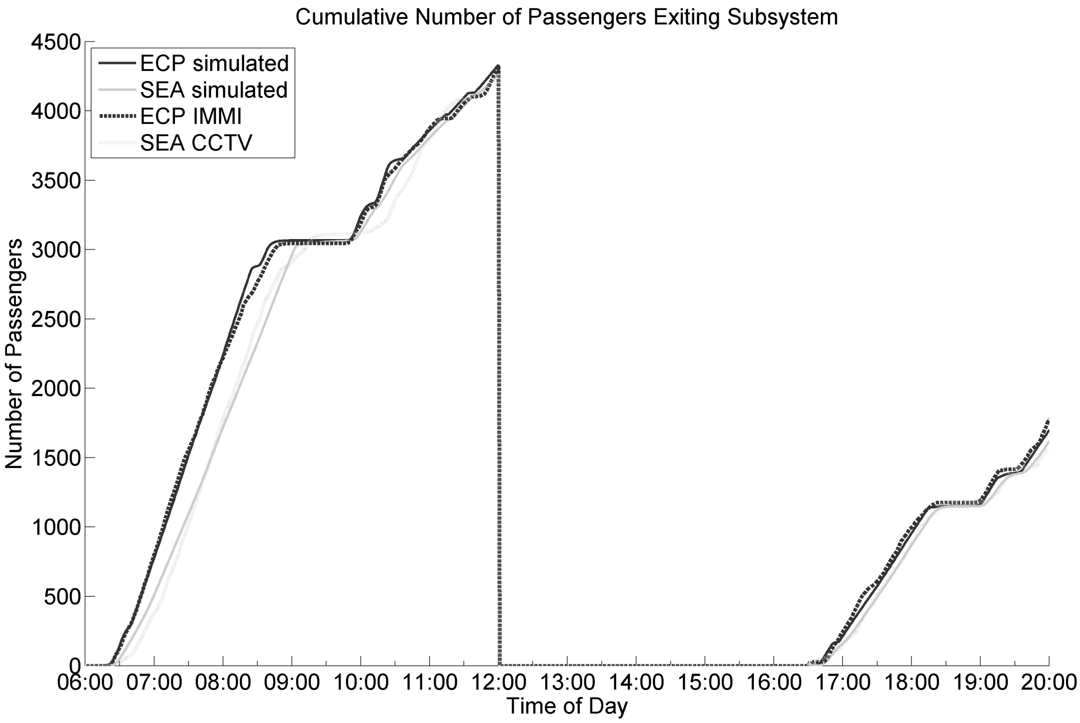

Figure 3: Model predicted cumulative exit curves for the ECP and SEA subsystems compared against the timestamp data derived ECP exit curve and CCTV derived SEA exit curve respectively.

Fig. 3 shows the predicted cumulative exit curves for the ECP and SEA compared to the timestamp derived exit curve for the ECP and CCTV derived exit curve for the SEA respectively. The dataset that was available for use was for 6:00-12:00 and 16:00-20:00 (a total of 10 hours) on Sunday, September 30 2012 at Brisbane International Airport. It can be seen that the simulated curves closely match the data.

Note that expert knowledge and demographic data were used to train the CPTs for the nodes associated with cycle time in the ECP subsystem (i.e. SmartGate and Manual Check cycle time); hence the timestamp dataset independently validates the results. Similarly, the CCTV derived SEA curve independently validates the model predictions for the SEA. Note that the SEA curve captures the accumulated effects over the entire inbound process as it is the final area in the facilitation process.

The root mean squared error, or RMSE, provides a measurement of absolute error in the simulated cumulative curve. RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left(\hat{X}(k) - X(k)\right)^2} \tag{11}$$

where $\hat{X}(k)$ is the model estimated value at time slice k and $X(k)$ is the value being benchmarked against for n time slices. The RMSE for the model simulated ECP and SEA exit curves, as well as the coefficient of determination R^2 value for goodness of fit are shown in Table 2.

	ECP Cumulative Curve	SEA Cumulative Curve
RMSE	42.2 Passengers	70.0 Passengers
R^2	0.9994	0.9982

Table 2: Goodness of fit results.

It can be seen that not only does the proposed model capture passenger flow, it also demonstrates strong predictive validity as shown with the cumulative curves, RMSE and R^2 values above. In addition, the model also passed a comprehensive validation framework put forward by Pitchforth and Mengersen (2013). This framework includes expert based face validity as performed with operational managers in airport industry and government partner organisations, extreme conditions testing (e.g. showing the model outputs zero when there are no flights), and comparisons against other airport models.

4.3. Discussion

The primary motivation for the HQBN is the ability to perform explicit, quantitative causal analysis with an extensible framework as discussed in Section 2. As shown in the preceding section, the model innately captures passenger flow through the queuing component. This section demonstrates a selection of applications of the IPFM to support airport decision making and especially highlights the combination of extensibility and causal analysis.

The model is able to simulate and analyse a wide range of operational scenarios such as that relating to different aircraft arrival scenarios, different resourcing configurations (e.g. number of staff rostered), and different passenger demographics (e.g. age, nationality). As an example, a peak traffic analysis (one of the main challenges facing airports (de Neufville and Odoni, 2003)) scenario is presented in Fig. 4 based on an assumed flight schedule and staff rostering schedule. There is a large morning peak at the ECP exceeding 1200 passengers, peaking at approximately 7:30am (see Fig. 4). By examining the throughput of each of the four consecutive areas in the inbound process (see Fig. 5), it can be seen that the throughput of the AC greatly exceeds that of the downstream ECP subsystem. As a result, this produces the peak seen in Fig. 4. Note further that the throughput of the ECP approximately matches that of the BH, hence, the peak in that area at approximately 9:00am is substantially smaller.

The preceding scenario focused on the passenger flow (i.e. queuing) aspect of the facilitation problem; consider a more complex scenario focusing on the SEA that involves causal analysis. In Fig. 4, it can be seen that the number of passengers in the SEA is increasing around 07:30. Consider a scenario where the operator(s) are interested in ascertaining the factors affecting passenger flow in the SEA. Note that due to security and commercial sensitivities, the numbers presented in this case study are for a demonstration scenario and do not reflect the actual values used for model validation.

As described in Section 3.2, passenger flow in the queuing model is linked to the BN via cycle times. In the case of the SEA, the throughput for these subsystems are linked to: (i) interview time, (ii) K9 (canine based) processing

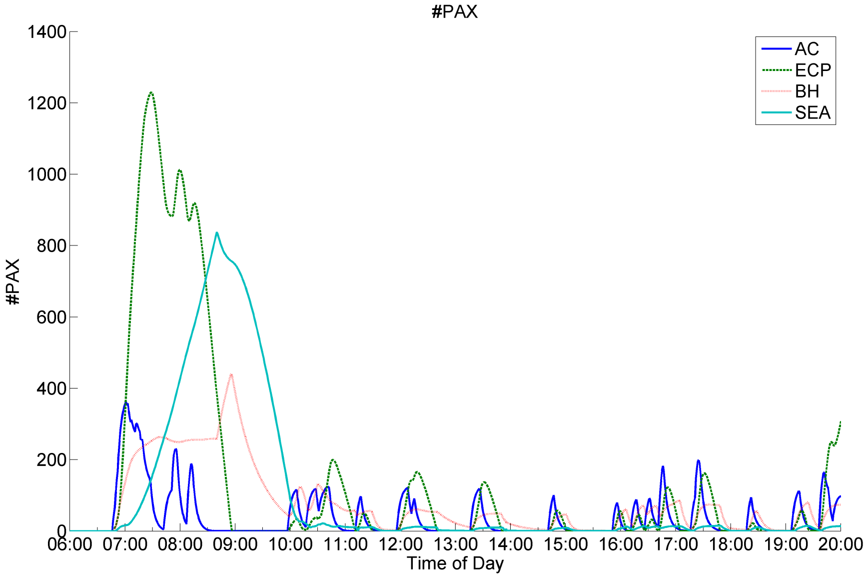

Figure 4: The number of passengers in each of the four main areas, showing a substantial morning peak at the ECP.

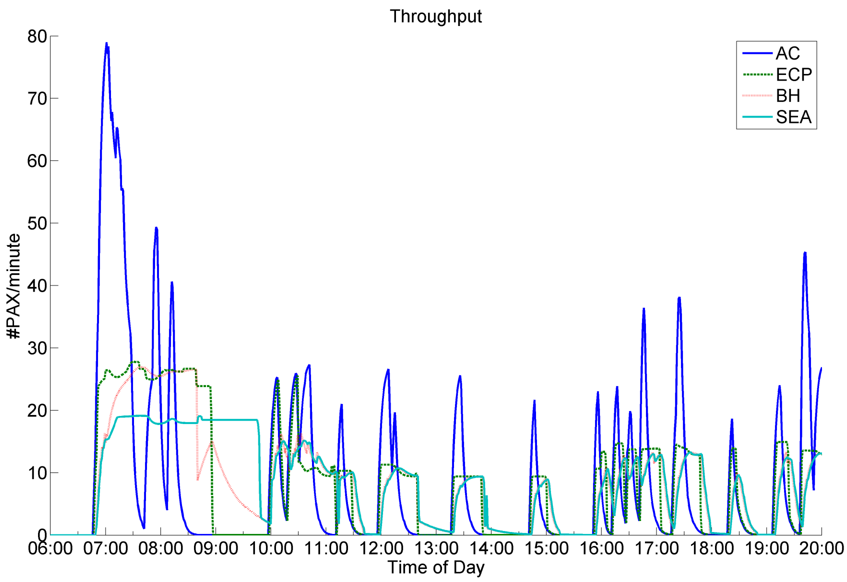

Figure 5: The throughput rate (passengers per minute) in each of the four main areas. Note the discrepancy between the upstream AC and downstream ECP throughputs.

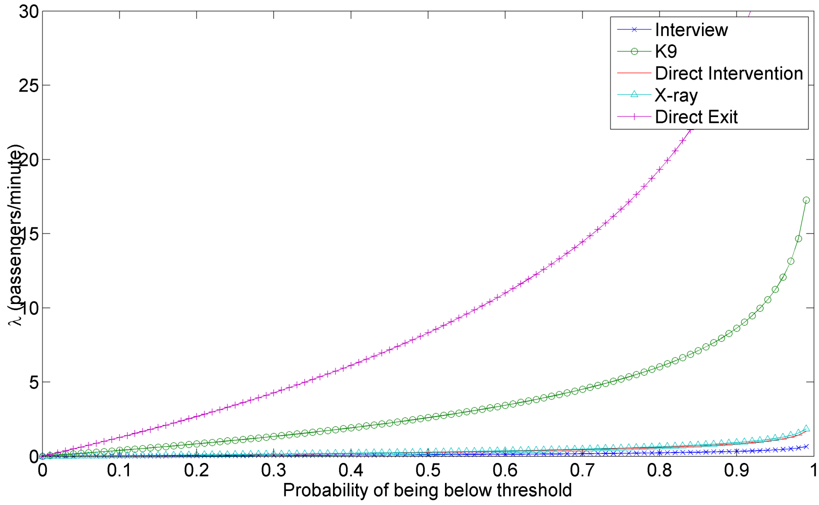

Figure 6: Passenger throughput rates plotted against probability of being below the threshold time using (5).

time, (iii) direct intervention (bench based) processing time, (iv) X-ray processing time, and (v) direct exit mean dwell time (small but non-zero time to leave queuing area). Note further that the discretised time threshold for direct exit is considerably smaller than that corresponding to the other sub-systems, which corresponds to larger λ (passenger throughput) values as illustrated in Fig. 6. Additionally, note that for this model, the overall throughput of the SEA is calculated as the sum of all five SEA subsystems. The passenger flow of each subsystem is a function of both the number of modules and λ as shown in (8) and the proportion of passengers who go to each subsystem α (refer to (10)); the latter of these is characterised by the marshal decisions (refer to Section 4.1). A subset of the factors influencing passenger flow is illustrated in Fig. 7; note that the values shown reflect those at 7:30am.

The BN provides a direct visual representation of the causal links between different factors via directed arcs as shown in Fig. 7. The number of passengers in the area is discretised into intervals (e.g. 1-10, 11-20 passengers etc.) whereas the time nodes, comprising SEA dwell time as well as the cycle time nodes (e.g. interview time, K9 processing time etc.), are discretised into binary states of the form above or below a threshold time. The two marshal decision nodes are used to determine α where ACBPS refers to the Customs and Border Protection marshal and DAFF is the Biosecurity marshal. Finally, the passenger risk profile, passenger complexity and overall biosecurity risk are shown as examples of some of the 'other factors' that can be captured in a BN. Note that the spatial aspect of the system is captured indirectly via time based nodes (e.g. time to walk through a space) and nodes about the number of passengers and congestion.

However, behind each of the nodes illustrated in Fig. 7 is a CPT, an example of which is provided in Table 3 for the DAFF Marshall Decision node. It can be seen that the probability of decision outcomes are quantitatively captured as a function of the parent nodes, namely, Passenger Complexity and Biosecurity Risk. As a result of this quantification, it is possible to analyse cause-effect relationships through: (i) forward reasoning and simulation, (ii)

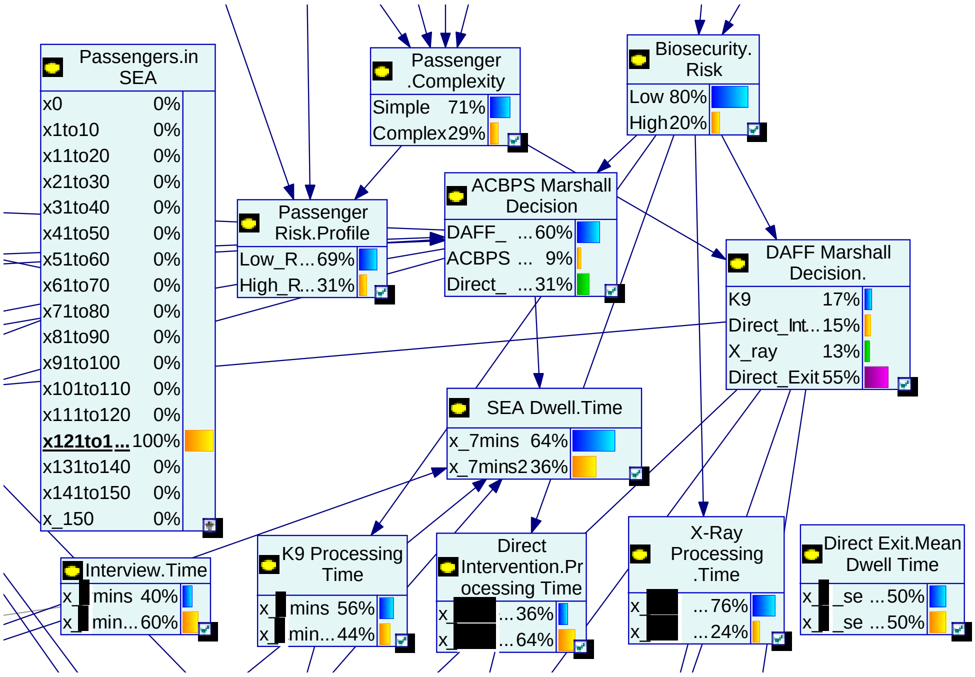

Figure 7: Extract from the SEA BN showing the relationship between factors such as biosecurity risk, the number of passengers in the SEA, and SEA dwell time and cycle time nodes. Note that there are many other links to other nodes outside of the subset shown.

Table 3: Example CPT for the DAFF Marshall Decision node whose parent nodes are Passenger Complexity and Biosecurity Risk (see Fig. 7). Note that due to security sensitivities, the numbers presented are for demonstration purposes only.

Passenger Complexity	Simple	Simple	Complex	Complex
Biosecurity Risk	Low	High	Low	High
K9	0.02	0.4	0.45	0.025
Direct Intervention	0.02	0.2	0.25	0.95
X-ray	0.04	0.4	0.25	0.025
Direct Exit	0.92	0	0.05	0

Table 4: Comparison of BN node state probabilities for the scenario where the biosecurity risk is high versus the original scenario in Fig. 7. Note that Proc. T refers to "Processing Time".

Above Threshold Probability	K9 Proc. T	Direct Intervention Proc. T	X-Ray Proc. T
Original scenario	50%	64%	24%
New scenario	50%	80%	40%

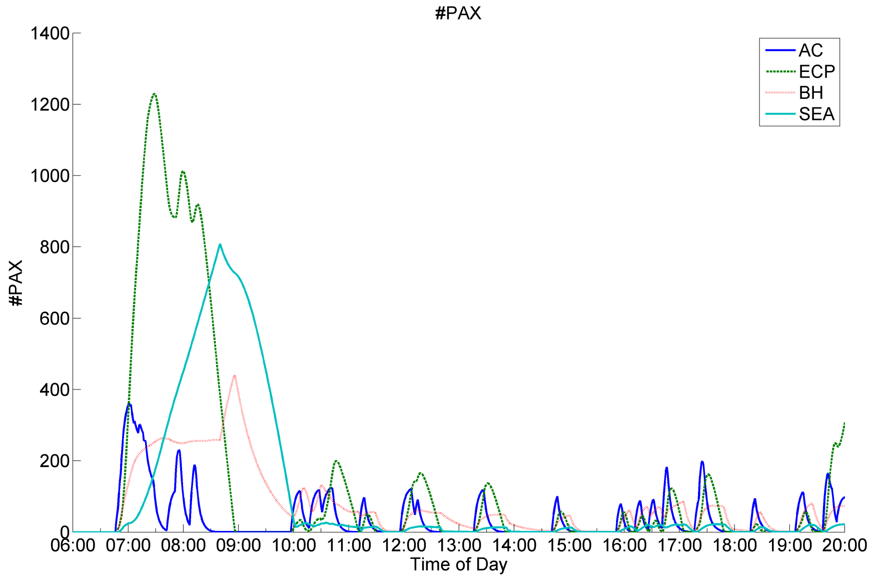

Figure 8: The number of passengers in each of the four main areas for a high biosecurity risk scenario; note the higher peak in the SEA compared to Fig. 4.

backwards reasoning and simulation, and (iii) sensitivity and strength of influence analysis.

Consider firstly the use of forward reasoning and simulation. If the biosecurity risk is high, it has an effect on various factors including processing time in the network as shown in Table 4, comparing processing time probabilities between the original scenario and the high biosecurity risk scenario. Note that in the CPT for K9 processing time, the conditional probabilities remain the same for low or high biosecurity risk, hence there is no change in processing time.

In addition, the processing time affects the λ value, as a result, the resultant effect on passenger flow is non-trivial, producing a peak at SEA of approximately 800 passengers as shown in Fig. 8 compared to 700 pasengers in Fig. 4.

Conversely, consider the scenario where the user wishes to understand what combination of factors would most likely result in a SEA dwell time of below 7 minutes. Known as backwards propagation or backwards inferencing, the results can also provide insight into identifying cause-effect relationships from a diagnostic perspective; the change in node state probabilities for the marshal decision nodes and the biosecurity risk node is shown in Table 5. It can be seen that such a scenario will likely involve a substantial increase in the number of direct exit passengers, as well as a

Table 5: Comparison of BN node state probabilities for the scenario where the SEA dwell time is below the threshold versus the original scenario in Fig. 7.

	ACBPS Marshall Decision			Biosecurity Risk
	DAFF	ACBPS	Direct Exit	Low
Original scenario	60%	9%	31%	80%
New scenario	51%	0%	48%	89%

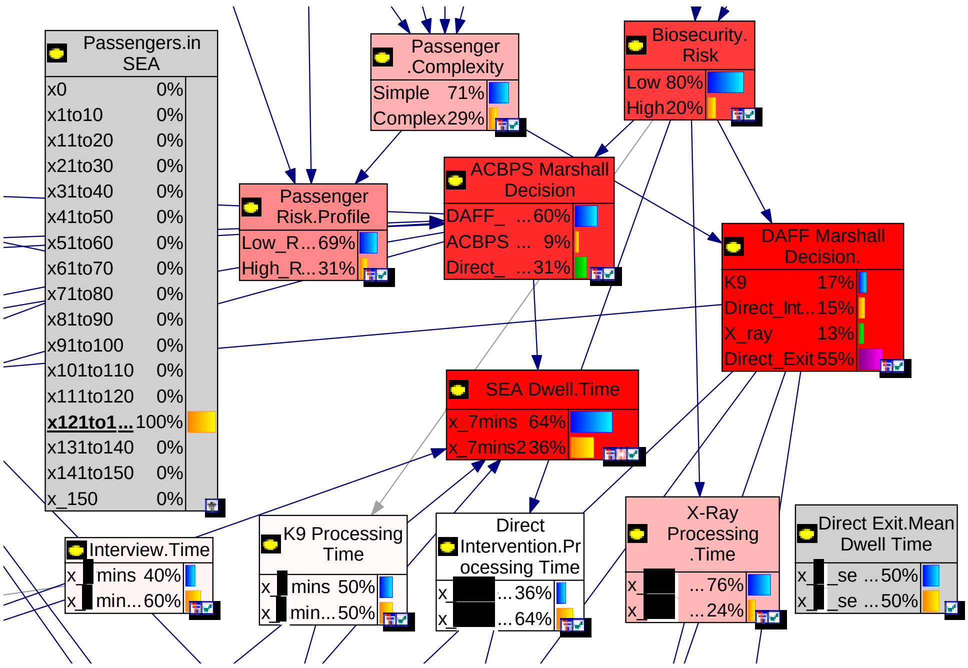

Figure 9: BN diagram showing the sensitivity of SEA Dwell Time to the other factors, where a darker shading of red represents greater sensitivity.

'lowering' of biosecurity risk. Note that, as before, the model is set up for the operational scenario at 7:30am.

Finally, consider again the 7:30am operational scenario and this time, the goal is to ascertain what are the main factors affecting SEA dwell time. By applying a sensitivity analysis, it is possible to show that dwell time is highly sensitive (shown in red in Fig. 9) to marshall decisions as well as biosecurity risk. The interplay between biosecurity risk and pasenger flow is intuitive in that higher risk leads to reduced passenger flow and lower risk leads to higher flow.

The above scenarios show three different ways to perform explicit, quantitative causal analysis using the proposed HQBN framework. It is able to capture a range of factors including biosecurity risk in addition to the core task of modelling passenger flow. In addition, it captures space through factors such as congestion and indirectly via time based factors, and enables a range of 'what-if' analyses. Therefore, it provides the ability to evaluate stakeholder KPIs, address the key requirements for passenger facilitation modelling and address the key gap in the literature as developed in Table 1. In addition, the approach has a number of other advantages: the model can be quantified with a combination of expert knowledge as well as existing sources of data (see Section 4.2), and can represent uncertainty through the use of probability distributions.

5. Conclusion

It can be seen that the airport terminal is a complex system with a wide variety of operational factors that can influence the flow of passengers, which in turn, can influence the performance of said factors. These factors include passenger demographics and risk factors and the interactions between them and individual flights, airport processes and passenger flows. It was identified, through consultation with expert end-users, that a decision support model must have the ability to capture passenger flow, incorporate the effects of space and enable 'what-if' analysis. In addition, it needs to be extensible to capture other operational factors relevant to airport operations such as demographics and risk, whilst providing an explicit, quantitative ability to identify the root cause of performance issues. Existing airport terminal passenger models do not provide simultaneously address the need for explicit, quantitative causal analysis and extensibility to model other operational factors.

The proposed framework exploits the explicit, causal analytic abilities of the BN, as well as its innate extensibility to capture other factors to address the key gap in the literature. By interfacing a BN with queuing theory, the proposed HQBN turns the static BN into a dynamic framework capable of simulating passenger flow and performing explicit, quantitative causal analysis. The framework achieves this by integrating a stochastic queuing model with the BN using the unique properties of the Poisson and Exponential distributions to transform between cycle or dwell time and passenger flow.

Such a framework for dynamic complex systems modelling could be generalised to arbitrary socio-technical systems where there is a flow of people and subsystems of processes or activities (Bostrom and Heinen, 1977). Future work includes the development of improved learning algorithms for learning BN CPTs in such a complex environment, exploration of other approaches to modelling mixtures of passengers and testing of the framework on different types of complex systems (such as hospitals).

6. Acknowledgements

This research forms part of the work undertaken by the project "Airports of the Future" (LP0990135) which is funded by the Australian Research Council Linkage Project scheme.

7. References

Appelt, S., Batta, R., Li, L., Drury, C., 9-12 Dec. 2007 2007. Simulation of passenger check-in at a medium-sized us airport.

Australian Competition and Consumer Commission (ACCC), 2013. Airport Monitoring Report 2011-12: Price, financial performance and quality of service monitoring. Canberra.

Bevilacqua, M., Ciarapica, F. E., 7-10 Dec. 2010 2010. Analysis of check-in procedure using simulation: A case study.

Bostrom, R. P., Heinen, J. S., 1977. Mis problems and failures: A socio-technical perspective. part i: The causes. MIS Quarterly 1 (3), 17–32.

Brunetta, L., Righi, L., Andreatta, G., 1999. An operations research model for the evaluation of an airport terminal: Slam (simple landside aggregate model). Journal of Air Transport Management 5 (3), 161–175.

Choy, S. L., O'Leary, R., Mengersen, K., 2009. Elicitation by design in ecology: Using expert opinion to inform priors for bayesian statistical models. Ecology 90 (1), 265–277.

de Neufville, R., Odoni, A. R., 2003. Airport systems planning design and management. McGraw-Hill.

Eilon, S., Mathewson, S., 1973. A simulation study for the design of an air terminal building. IEEE Transactions on Systems, Man and Cybernetics 3 (4), 308–317.

Gongora, M., Ashfaq, W., 0-0 0 2006. Analysis of passenger movement at birmingham international airport using evolutionary techniques.

Heckerman, D., Mamdani, A., Wellman, M. P., 1995. Real-world applications of bayesian networks. Communications of the ACM 38 (3), 24–26.

Helbing, D., Molnar, P., 1995. Social force model for pedestrian dynamics. Physical Review E 51 (5), 4282–4286.

International Civil Aviation Organisation, July, 2005. Annex 9 to the convention on international civil aviation: Facilitation. Tech. rep., ICAO.

Jim, H. K., Chang, Z. Y., 1998. An airport passenger terminal simulator: A planning and design tool. Simulation Practice and Theory 6 (4), 387–396.

Johnson, S., Mengersen, K., Waal, A. d., Marnewick, K., Cilliers, D., Houser, A. M., Boast, L., 2010. Modelling cheetah relocation success in southern africa using an iterative bayesian network development cycle. Ecological Modelling 221, 641–651.

Kiran, A. S., Cetinkaya, T., Og, S., 2000. Simulation modeling and analysis of a new international terminal.

Kjrulff, U., 1995. dhugin: a computational system for dynamic time-sliced bayesian networks. International Journal of Forecasting 11 (1), 89–111.

Kleinschmidt, T., Guo, X., Ma, W., Yarlagadda, P. K., 2011. Including airport duty-free shopping in arriving passenger simulation and the opportunities this presents.

Koch, D. B., 2004. 3d visualization to support airport security operations. IEEE Aerospace and Electronic Systems Magazine 19 (6), 23–28.

Lui, R., Nanda, R., Browne, J. J., 1972. International passenger and baggage processing at john f. kennedy international airport. IEEE Transactions on Systems, Man and Cybernetics 2 (2), 221–225.

Manataki, I. E., Zografos, K. G., 2009. A generic system dynamics based tool for airport terminal performance analysis. Transportation Research Part C: Emerging Technologies 17 (4), 428–443.

Manataki, I. E., Zografos, K. G., 2010. Assessing airport terminal performance using a system dynamics model. Journal of Air Transport Management 16 (2), 86–93.

Neapolitan, R. E., 2004. Learning Bayesian Networks. Pearson Prentice Hall, Upper Saddle River.

Newell, G. F., 1971. Application of Queuing Theory. Chapman and Hall, London.

Nombela, G., de Rus, G., Betancor, O., 2004. Internalizing airport congestion. Utilities Policy 12 (4), 323–331.

Odoni, A. R., de Neufville, R., 1992. Passenger terminal design. Transportation Research Part A: Policy and Practice 26 (1), 27–35.

Passenger Facilitation Taskforce, 2009. International Airport Operator's Guide, 1st Edition. Australian Customs and Border Protection Service.

Pearl, J., 1988. Probabilistic Reasoning in Intelligent Systems. Morgan Kaufmann.

Pitchforth, J., Mengersen, K., 2013. A proposed validation framework for expert elicited bayesian networks. Expert Systems with Applications 40, 162–167.

Popovic, V., Kraal, B., Kirk, P., 2009. Passenger experience in an airport: An activity-centred approach.

Pourret, O., Naim, P., Marcot, B., 2008. Bayesian Networks A Practical Guide to Applications. Wiley and Sons, West Sussex.

Ross, S. M., 2010. Introduction to probability models, 10th Edition. Academic Press, Amsterdam.

Schultz, M., Fricke, H., 2011. Managing passenger handling at airport terminals.

Takakuwa, S., Oyama, T., 2003. Modeling people flow: simulation analysis of international-departure passenger flows in an airport terminal.

Tosic, V., 1992. A review of airport passenger terminal operations analysis and modelling. Transportation Research Part A: Policy and Practice 26 (1), 3–26.

Wilson, D., Roe, E. K., So, S. A., 3-6 Dec. 2006 2006. Security checkpoint optimizer (sco): An application for simulating the operations of airport security checkpoints.

Wu, P. P.-Y., Mengersen, K., 2013. A review of models and model usage scenarios for an airport complex system. Transportation Research Part A: Policy and Practice 47, 124–140.

Yanagisawa, D., Suma, Y., Tomoeda, A., Miura, A., Ohtsuka, K., Nishinari, K., 2013. Walking-distance introduced queueing model for pedestrian queueing system: Theoretical analysis and experimental verification. Transportation Research Part C: Emerging Technologies 37 (0), 238–259.

Zografos, K. G., Madas, M. A., 2006. Development and demonstration of an integrated decision support system for airport performance analysis. Transportation Research Part C: Emerging Technologies 14 (1), 1–17.